U.S. ATLAS Computing Facilities

Rich Baker

Brookhaven National Laboratory

US ATLAS Computing Meeting

BNL, August 27-29, 2003

BROOKHAVEN NATIONAL LABORATORY

Mission of US ATLAS Computing Facilities

- Supply capacities to the ***ATLAS Distributed Virtual Offline Computing Center***
  - At levels agreed to in a computing resource MoU (Yet to be written)
- Guarantee the Computing Required for Effective Participation by U.S. Physicists in the ATLAS Physics Program
  - Direct access to and analysis of physics data sets
  - Simulation, re-reconstruction, and reorganization of data as required to support such analyses

ATLAS Facilities Model

- ATLAS Computing Will Employ the ATLAS Virtual Offline Computing Facility to process and analyze its data
  - Distributed set of resources including:
    - CERN Tier 0
    - All Regional Facilities (Tier 1's) - Typically ~200 users each
    - Some National Facilities (Tier 2's)
  - All members of ATLAS Virtual Organization (VO) must contribute in funds or in kind (personnel, equipment), proportional to author count
  - All members of ATLAS VO will have defined access rights
- Typically only a subset of resources at a regional or national center are Integrated into the Virtual Facility
  - Non-integrated portion over which regional control is retained is expected to be used to augment resources supporting analyses of region interest

Analysis Model: All ESD Resident on Disk

- Enables ~24 hour selection/regeneration passes (versus ~month if tape stored) – faster, better tuned, more consistent selection
- Allows navigation for individual events (to all processed, *though not Raw,* data) without recourse to tape and associated delay – faster more detailed analysis of larger consistently selected data sets
- Avoids contention between analyses over ESD disk space and the need to develop complex algorithms to optimize management of that space – better result with less effort
- **Complete set on disk at US Tier 1**
  - **Reduced sensitivity to performance of multiple Tier 1's, intervening network (transatlantic) & middleware – improved system reliability, availability, robustness and performance – cost impact discussed later**

US ATLAS Facilities

- A Coordinated Grid of Distributed Resources Including …
  - Tier 1 Facility at Brookhaven – Rich Baker / Bruce Gibbard
    - Currently operational at ~1% of required 2008 capacity
  - 5 Permanent Tier 2 Facilities – Saul Youssef
    - Scheduled for selection beginning in 2004
    - Currently there are 2 Prototype Tier 2's
      - **Indiana U – Fred Luehring / University of Chicago – Rob Gardner**
      - **Boston U – Saul Youssef**
  - 7 Currently Active Tier 3 (Institutional) Facilities
  - WAN Coordination Activity – Shawn McKee
  - Program of Grid R&D Activities – Rob Gardner
    - **Based on Grid Projects (**PPDG, GriPhyN, iVDGL, EU Data Grid, EGEE, etc.**)**
  - Grid Production & Production Support Effort – Kaushik De/Pavel Nevski

BNL Tier 1 Facility

- Functions
  - Primary U.S. data repository for ATLAS
  - Programmatic event selection and AOD & DPD regeneration from ESD
  - *Chaotic* high level analysis by individuals
    - Especially for large data set analyses
  - Significant source of Monte Carlo
  - Re-reconstruction as needed
  - Technical support for smaller US computing resource centers
- Co-located and operated with the RHIC Computing Facility
  - To date a very synergistic relationship
  - Some recent increased divergence
  - Substantial benefit from cross use of idle resources (2000 CPU's)

Tier 1 Facility Current Deployment

- ~60 Dual Processor Linux Nodes
  - 16 Available for Interactive Login
  - Limited Temporary Local Disk Space /home/tmp/
- ~10 TB NFS Disk Space – Visible from All Nodes
  - 250 GB Home Directories: /usatlas/u/, Initial 500 MB Quota
  - 500 GB Work Area: /usatlas/workarea/
  - 870 GB Scratch: /usatlas/scratch/
- ~500 GB AFS Disk Space – Accessible Worldwide
  - User Directories: /afs/usatlas/users/, Initial 200 MB Quota
- HPSS Tape System
- LSF and CONDOR Batch Systems

Facility Web Pages

- Home Page: http://www.acf.bnl.gov/
  - Note Link – New Users: Getting Started Guide
  - Also – Submit Problem Report
- Primary US ATLAS Web Server www.usatlas.bnl.gov
- User Pages Can Be Created
  - In Your AFS Area, Create: /afs/usatlas/users/username/WWW/
  - This Directory is Visible as: http://www.usatlas.bnl.gov/~username/
  - 200 MB Initial Quota on Your AFS Area

Tier 1 Facility Evolution for FY ‘04

- Modest equipment upgrades planned for FY ’04 (for DC 2)
  - Disk: 12 TBytes → 25 TBytes (factor of 2)
  - CPU Farm: 30 kSPECint2000 → 130 kSPECint2000 (factor of 4)
    - First processor farm upgrade since FY ’01 (3 years)
  - Robotic Tape Storage: 30 MBytes/sec → 60 MBytes/sec (factor of 2)

Capital Equipment

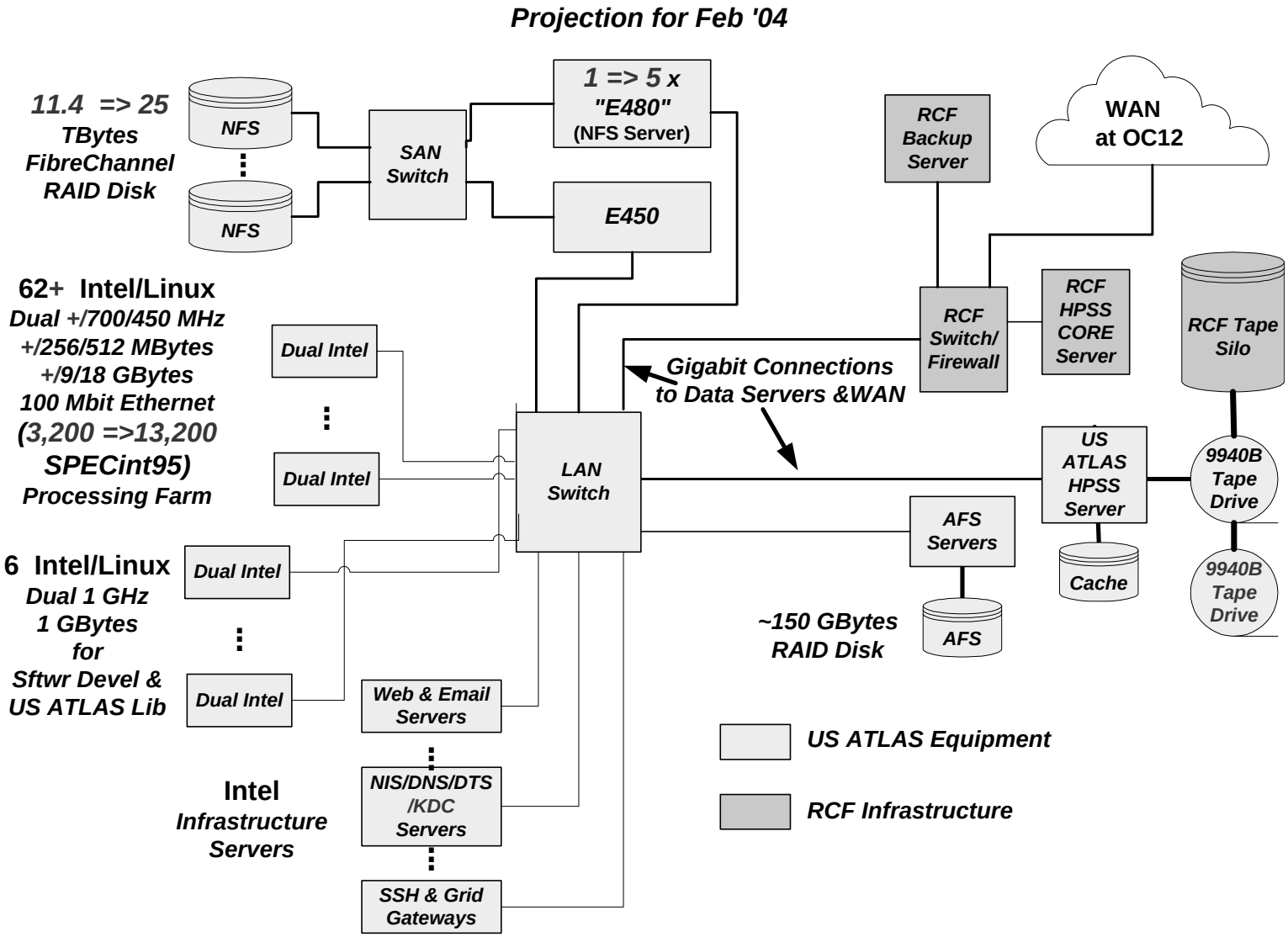
Tier 1 Capacity Profile

	2001	2002	2003	2004	2005	2006	2007	2008
CPU (kSPECint2000)	30	30	30	130	250	750	1,500	5,000
Disk (TBytes)	0.5	12	12	25	50	143	300	1,000
Disk (MBytes/sec)	40	90	90	500	1,000	3,000	6,000	20,000
Tape (PBytes)	0.01	0.05	0.05	0.10	0.21	0.32	0.86	2.05
Tape (MBytes/sec)	10	30	30	60	60	120	240	360
WAN (Mbits/sec)	155	155	622	622	2488	2488	9952	9952

Tier 1 Captial Equipment Cost Profile (At Year $k)

	2001	2002	2003	2004	2005	2006	2007	2008
CPU	$ 30	$ -	$ -	$ 123	$ 112	$ 282	$ 279	$ 881
Disk	$ 100	$ 137	$ -	$ 186	$ 201	$ 531	$ 544	$ 1,619
Tertiary Storage	$ 46	$ 25	$ -	$ 30	$ 170	$ 30	$ 80	$ 30
LAN	$ 79	$ -	$ 20	$ 20	$ 90	$ 100	$ 250	$ 250
Overhead	*$ 22*	*$ 14*	*$ 2*	*$ 32*	*$ 50*	*$ 83*	*$ 101*	*$ 245*
Total	**$ 277**	**$ 176**	**$ 22**	**$ 391**	**$ 624**	**$ 1,026**	**$ 1,254**	**$ 3,025**

US ATLAS Regional Center (Tier 1) at BNL



2.3.2 Tier 2 Facilities

- 5 Permanent Tier 2 Facilities
  - Primary resource for simulation
  - Empower individual institutions and small groups to do autonomous analyses using more directly accessible and locally managed resources
- 2 Prototype Tier 2's selected for ability to rapidly contribute to Grid development
  - Indiana University / (*effective FY '03*) University of Chicago
  - Boston University
- Permanent Tier 2 will be selected to leverage strong institutional resources
  - Selection of first two scheduled for spring 2004
  - Currently 7 active Tier 3's in addition to prototype Tier 2's; all candidates Tier 2's
- Aggregate of 5 permanent Tier 2's will be comparable to Tier 1 in CPU

Prototype Tier 2 Facilities

Indiana/Chicago Tier 2

IU

- 17 PII cluster decommissioned
- 64 Xeon 2.4 Ghz dedicated (new)
- 384 Xeon 2.4 Ghz shared (new)
- Lan: Myranet,Gigabit, Wan: OC12
- 1.5 T 10,000 rpm disks.

UC

- 4+64 Xeon 2.66 Ghz
- Lan: GigE switch, Wan, OC12
- 4 Tbyte disk

Boston University Tier 2

- 16 PIII dedicated
- 64 PIII shared
- 48 Xeon 2.4 Ghz (being set up)
- 32 Xeon 2.8 Ghz (being ordered)
- + Major CPU/disk purchase just before DC2
- Lan: Gigabit, Wan: OC12
- 1 Tbyte disk

Current software environments: RH 7.3**,** Atlas Production, VDT 1.1.10, Grid3

Current use: DC1 production as part of the U.S. Atlas testbed, DTG/Grid3 grid development. New users are welcome.